How accurate is my hotspot map?

Spencer Chainey
Director of Geographical Information Science
The Jill Dando Institute of Crime Science
University College London

UCL Jill Dando Institute of Crime Science

UCL Jill Dando Institute of Crime Science

Introduction

- Reasons for the research
- What we tested
- What we found out
- Some things to think about

Reasons for the research

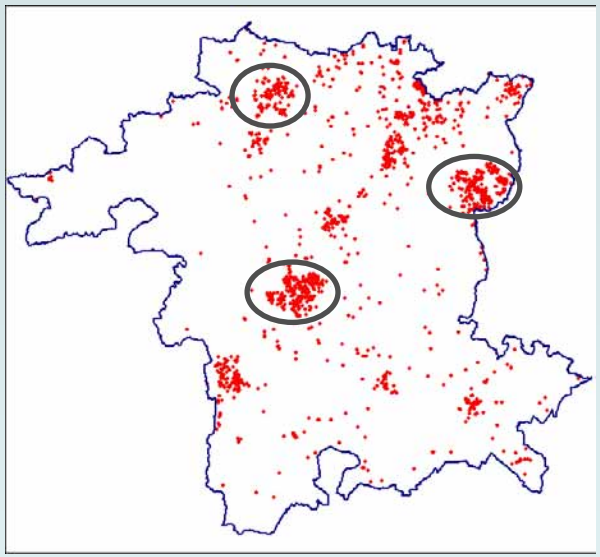
- What do we use hotspot maps for?
  - Identifying concentrations of crime, targeting/deploying resources and monitoring their impact
  - Create hotspot maps using data from the past to direct our future actions
  - Hotspot mapping is the most basic form of prediction

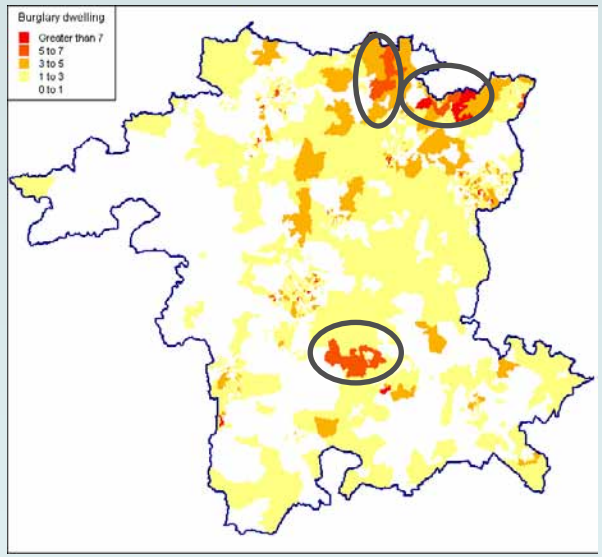
- Reviews of hotspot mapping techniques
  - Chainey, S.P. and Ratcliffe, J.H. (2005). GIS and Crime Mapping. John Wiley and Sons: Chichester
  - Eck. J., Chainey, S.P., Cameron, J., Leitner, M. and Wilson, R. (2005). Mapping crime: Understanding hotspots. Washington DC: National Institute of Justice.
  - Chainey, S.P., Reid, S., and Stuart, N. (2002). 'When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime' in 'Innovations in GIS 9' edited by Higgs, G. Taylor and Francis, London.
  - Jefferis, E. (1999) A multi-method exploration of crime hot-spots: a summary of findings. Crime Mapping Research Centre intramural project, The National Institute of Justice, Washington D.C.
  - Ratcliffe, J. and McCullagh, M.J. (1998). Hotbeds of crime and the search for spatial accuracy. Paper presented to the Second Crime Mapping Research Center Conference: Mapping Out Crime, Arlington, Virginia, USA. December 10–12, 1998.

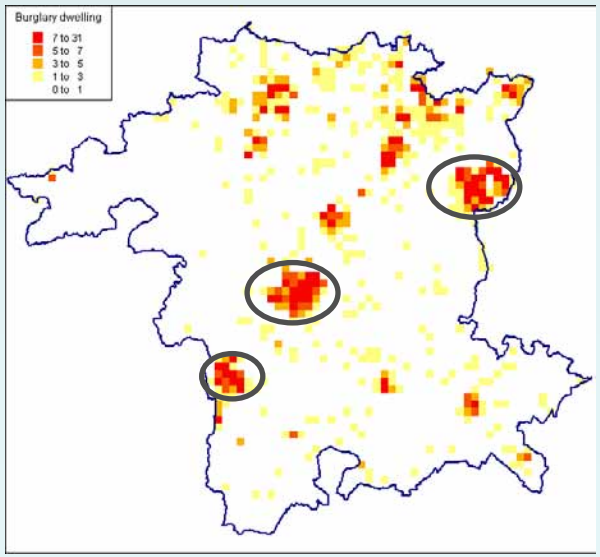
Reasons for the research

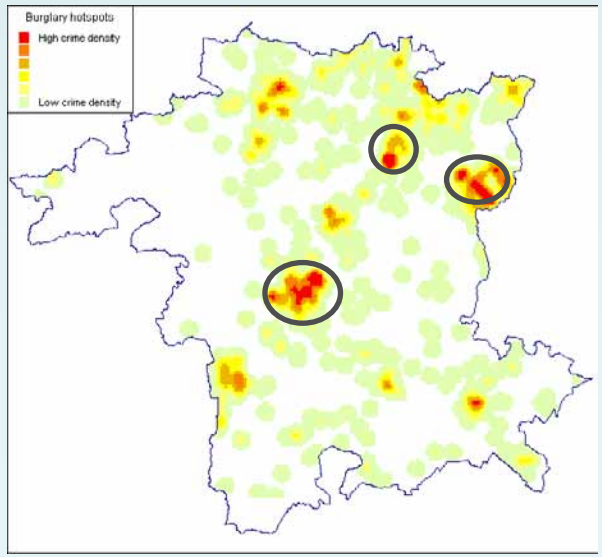
- Findings from reviews of hotspot mapping techniques
  - Use of preliminary statistical tests: nearest neighbour index and standard distance useful tests for clustering and dispersion
    - NNI useful for showing if you have enough data to create a hotspot map
  - Different hotspot mapping techniques produce different results
    - Shown differences between location, size and shape of areas that are defined as hotspots

Where are my three main hotspots?









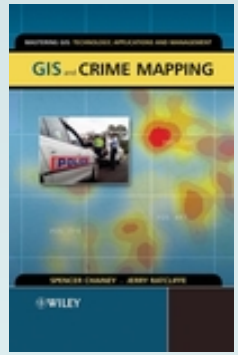
- Findings from reviews of hotspot mapping techniques
  - Some methods are easier to use than others
  - Users prefer the look of certain hotspot maps over others
- Have not been much more than a beauty contest between techniques

Reasons for the research

- What would be more useful …
  - Is my hotspot map any good at telling me what happens next
    - where will crime happen next?
    - where should I target my future actions?

Reasons for the research

- Prediction techniques: being developed, but providing a mixture of results
  - Univariate techniques
  - Leading indicators
  - Point process modelling
  - Artificial neural networks
  - Prospective mapping



For an introduction to these techniques see:

Chainey, S.P. and Ratcliffe, J.H. (2005). GIS and Crime Mapping. John Wiley and Sons: Chichester

- Prediction techniques: some way off from being a common tool on the crime analyst's desk
- What's the best 'common' hotspot mapping technique to use?

What we tested

- Which hotspot mapping technique is the best at telling me what happens next?
- Should I give my hotspot map shelf-life or use-by date?
  - Is my hotspot map only accurate for immediate use rather than being something that I can confidently refer to for many months?
- How does the currency of data influence the accuracy of my hotspot map?
  - Do I need to use 12 months worth of data or will last week's be sufficient?
- Does hotspot map accuracy vary by crime type?
  - Are my residential burglary hotspot maps just as good as the maps I produce for other crime types?

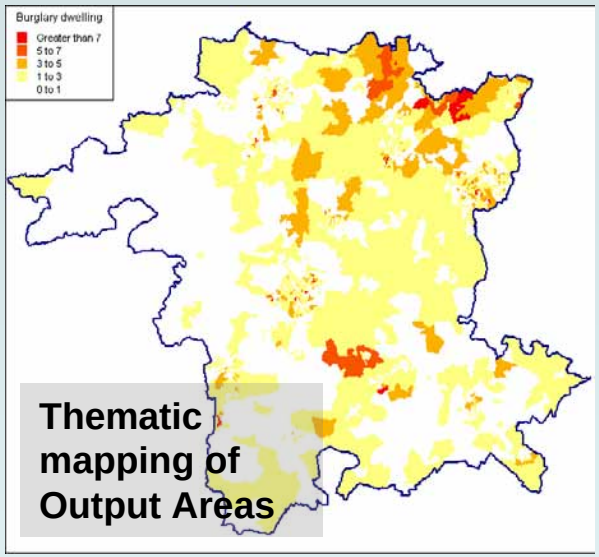
What we tested

Crime types and techniques

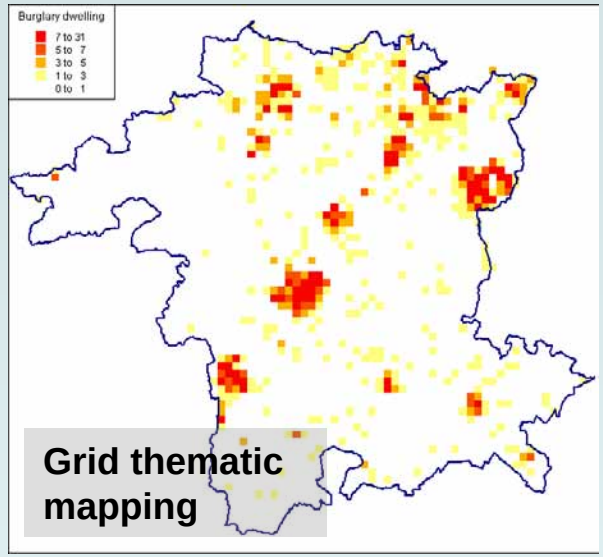
- Variety of crime types
  - Street crime (Robbery to the person and theft of the person (snatch theft))
  - Residential burglary (burglaries to domestic properties)
  - Theft of motor vehicles
  - Theft from motor vehicles
- Techniques
  - Choropleth mapping/thematic mapping of administrative boundary areas (e.g. Census Output Areas)
  - Spatial ellipses: standard deviational ellipses
  - Grid thematic mapping (uses a uniform grid)
  - Kernel density estimation (KDE)

UCL Jill Dando Institute of Crime Science

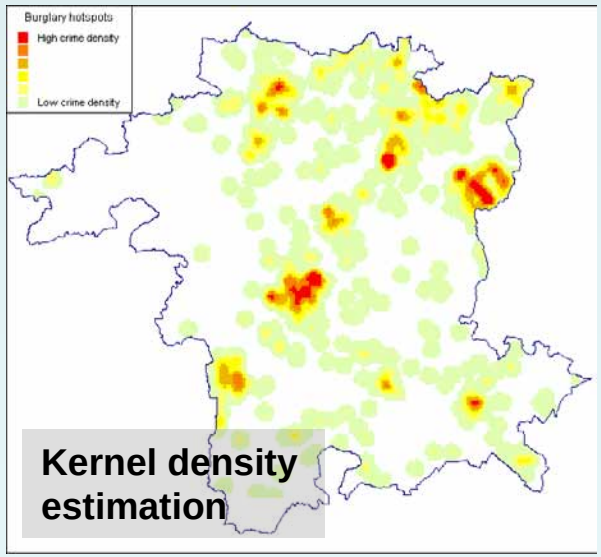
Hotspot mapping techniques



Thematic mapping of Output Areas



Grid thematic mapping



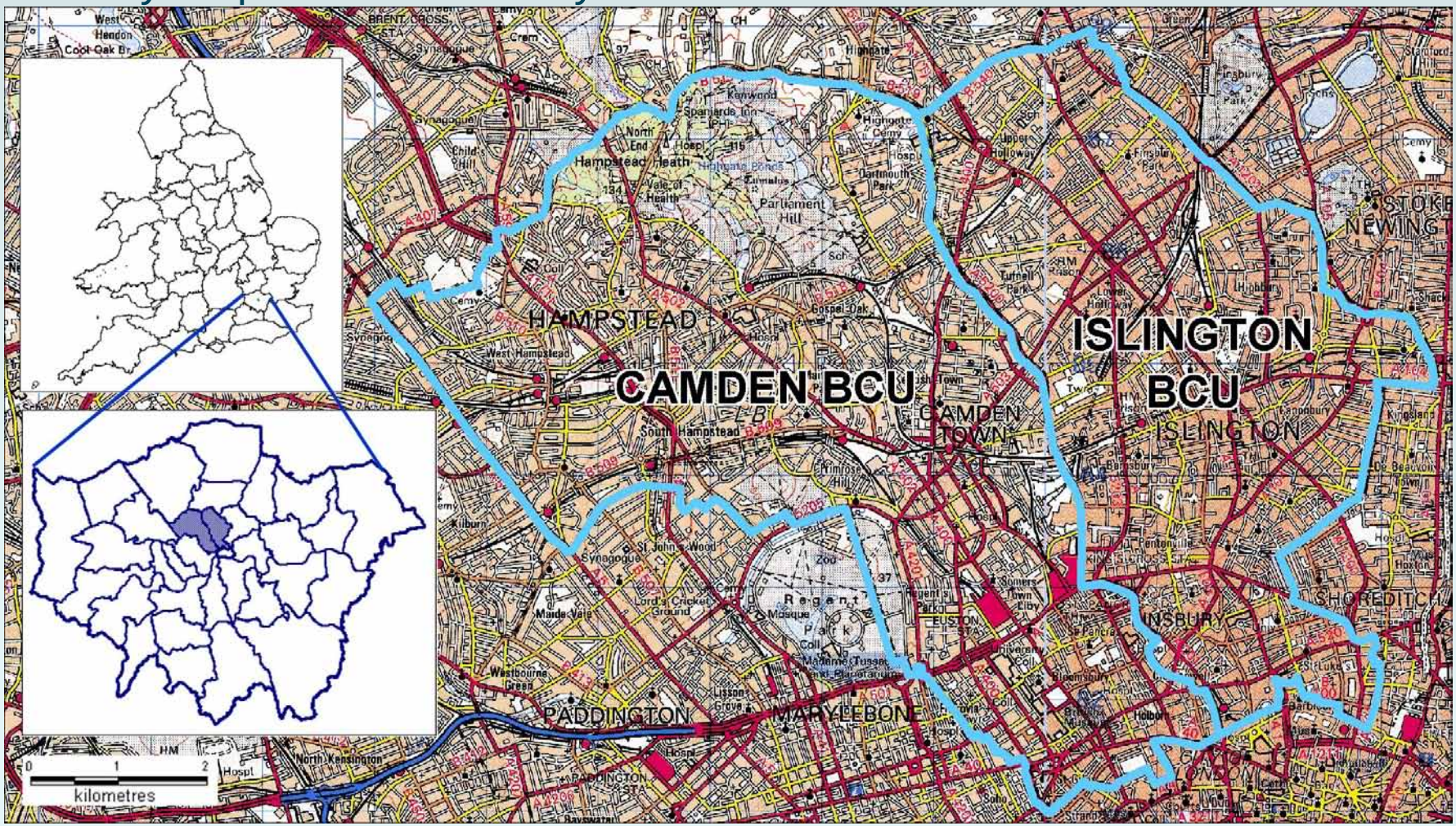
Kernel density estimation

UCL Jill Dando Institute of Crime Science

Study area

London Metropolitan Police: Camden and Islington BCUs

Two year period: 1st January 2002 to 31st December 2003



What we tested

Crime data

- Good quality data: Extensive data cleaning, geocoding and validation
- Split into two, 1-year data files
  - 2002: Input data (*training data*) – to create hotspot maps
  - 2003: Measurement data – to measure accuracy of hotspot maps in terms of what happens *next*

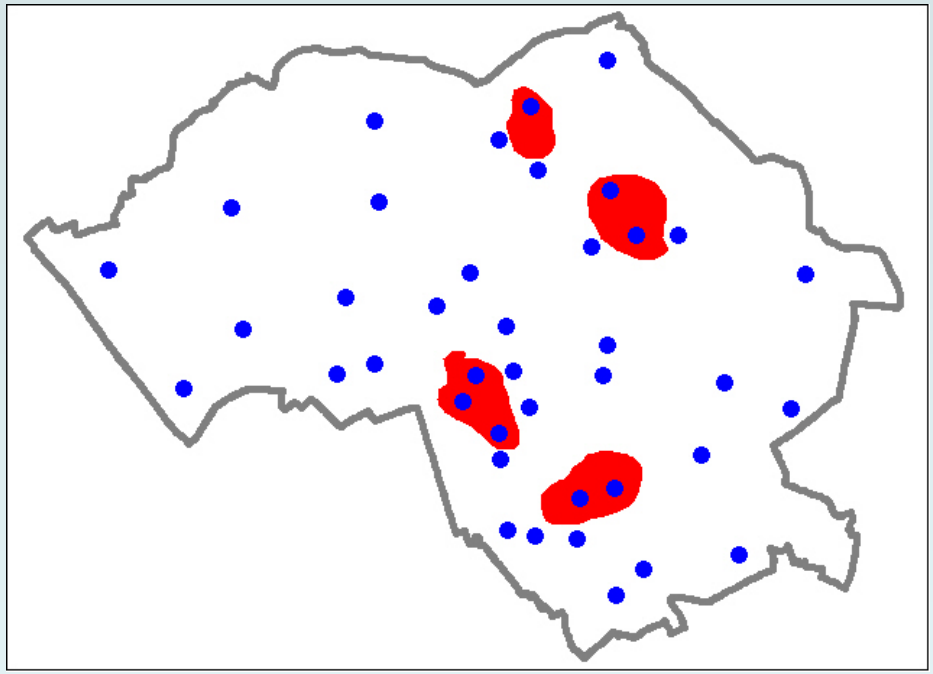
Number of crimes	Residential burglary	Street crime	Theft FROM vehicle	Theft OF vehicle
2002	6300	5249	10792	4243
2003	5671	4911	11536	4142

- Data cuts for Input and Measurement data
  - 24 hours, 2 days, 3 days, 1 week, 2 weeks, 1 month, 2 months, 3 months, 6 months, 12 months

2002 time periods used to create hotspot maps									
24 hours	2 days	3 days	1 week	2 weeks	1 month	2 months	3 months	6 months	12 months
31 Dec 02	30 Dec 02 - 31 Dec 02	29 Dec 02 - 31 Dec 02	25 Dec 02 - 31 Dec 02	18 Dec 02 - 31 Dec 02	01 Dec 02 - 31 Dec 02	01 Nov 02 - 31 Dec 02	01 Oct 02 - 31 Dec 02	01 Jul 02 - 31 Dec 02	01 Jan 02 - 31 Dec 02

Measuring hotspot map accuracy?

- Need to consider
  - New crimes that have occurred in identified hotspots
  - Area the hotspots cover relative to entire study area
- Measures
  - Hotspot hit rate: % of new crimes in hotspots
  - Search Efficiency Rate: crimes per km² in hotspots (Bowers et al, 2004)
  - Hotspot area: % area of hotspots relative to study area
  - Hotspot Accuracy Index (HAI): Hotspot hit rate/Hotspot area
- Wanted to keep it simple so others can also easily use it



Number of crimes: 40 Crimes in hotspots: 8
Hotspot hit rate: 20%

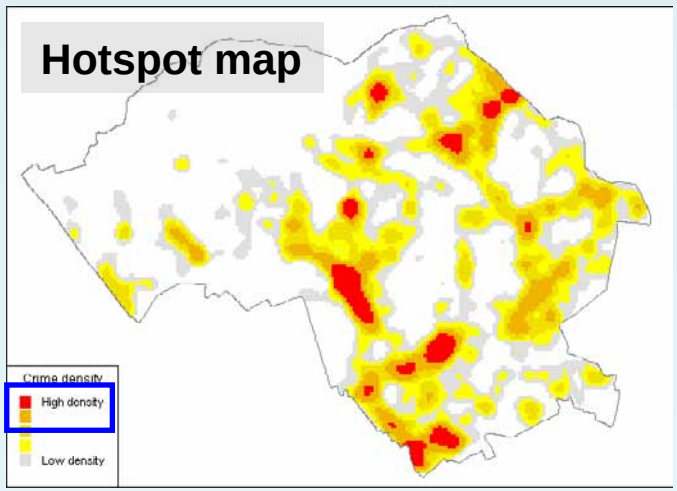
Hotspots coverage: 5 sq km
Search Efficiency Rate: 1.6

Study area: 50 sq km Hotspot area: 10%
Hotspot Accuracy Index: 2

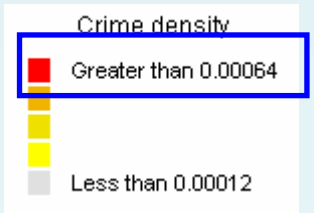
Measuring hotspot map accuracy?

- What areas do we determine as ‘hot’?
  - Ellipses: ellipses drawn
  - Thematic mapping approaches
    - Applied same thematic range method
      - Five thematic classes
      - Defaults from the quantile range method (MapInfo)
    - ‘Hot’ determined by the top thematic class

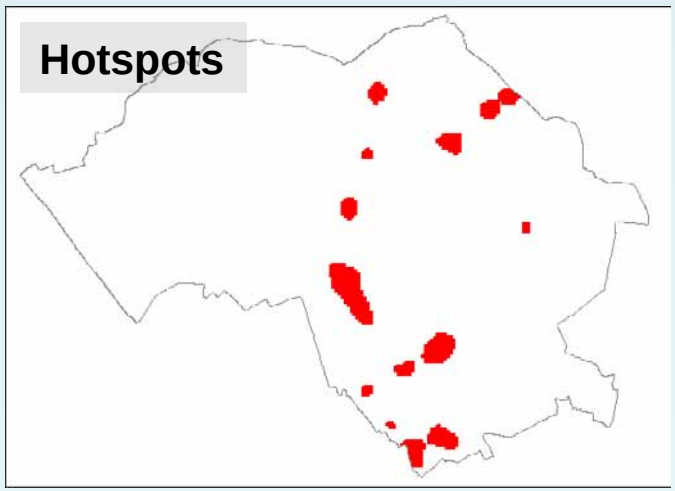
Hotspot map



Top thematic class identifies hotspots



Hotspots



Mapping techniques – parameters

- Thematic mapping of Output Areas (TMOA)
  - Approx 110 households in each OA
  - Mean OA size for Camden and Islington: 0.026 km^2
- Grid thematic mapping (GT)
  - 250m grid cell
- Kernel density estimation (KDE)
  - Cell size and bandwidth defaults in Hotspot Detective for MapInfo
  - Quartic method
- Spatial Ellipses (STAC)
  - STAC (CrimeStat): Standard deviational ellipses
  - Scan type: Triangular scan type
  - Study area boundary: Determined by the input crime dataset
  - Number of standard deviations for creation of each ellipse: 1
  - Search radius: 500m, 250m, Hotspot Detective default bandwidth
  - Minimum number of points per hotspot: Variable
- If I was a crime analyst and was going to use one of the techniques, these are the parameters I would tend to use

What we found out

Mapping techniques, crime type and hotspot accuracy

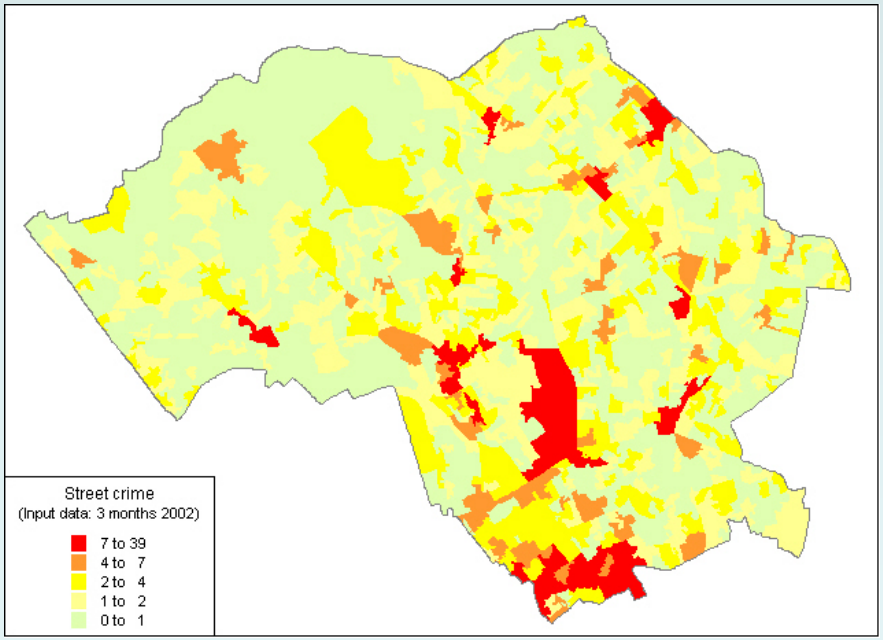
- KDE consistently produced high hotspot accuracy measures
- Standard deviational ellipses with 500m search radius were consistently the worst
  - but research showed influence of parameters on hotspot mapping accuracy
- Hotspot mapping accuracy differed between crime types

HAI values for 3 months 2002 data				
Mapping technique and crime type	Residential burglary	Street crime	Theft FROM vehicle	Theft OF vehicle
STAC-500	1.4	1.9	1.5	1.1
STAC-250	1.5	4.2	3.1	2.5
STAC-HD	1.9	3.3	3.5	1.2
TMOA	1.3	3.7	1.7	1.6
GT-250	1.6	4.9	2.5	1.9
GT-HD	1.8	4.5	2.3	2.2
KDE	2.2	5.8	3	2.2
n	*1564*	*1297*	*2851*	*942*

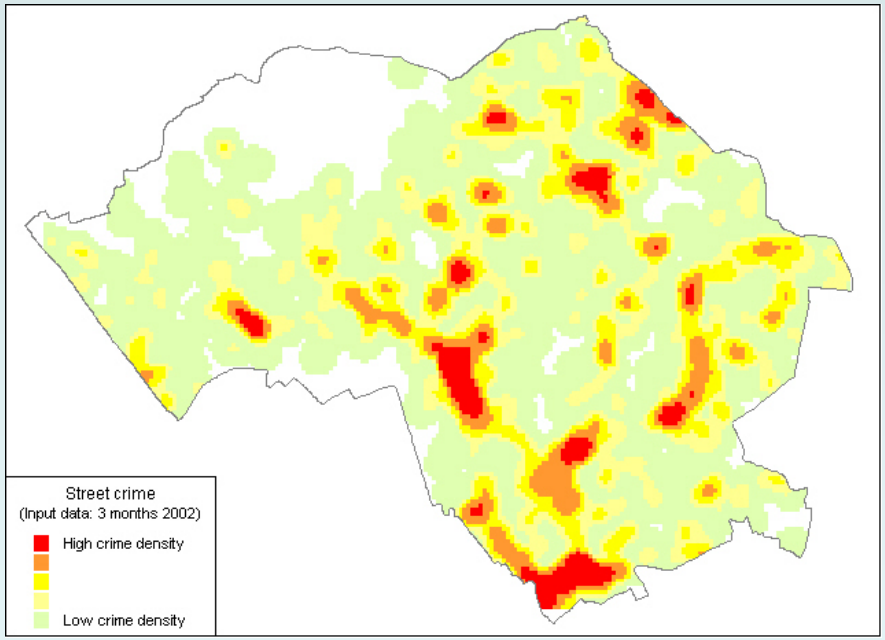
What we found out

Mapping techniques and hotspot accuracy

Thematic mapping of OAs HAI: 3.7



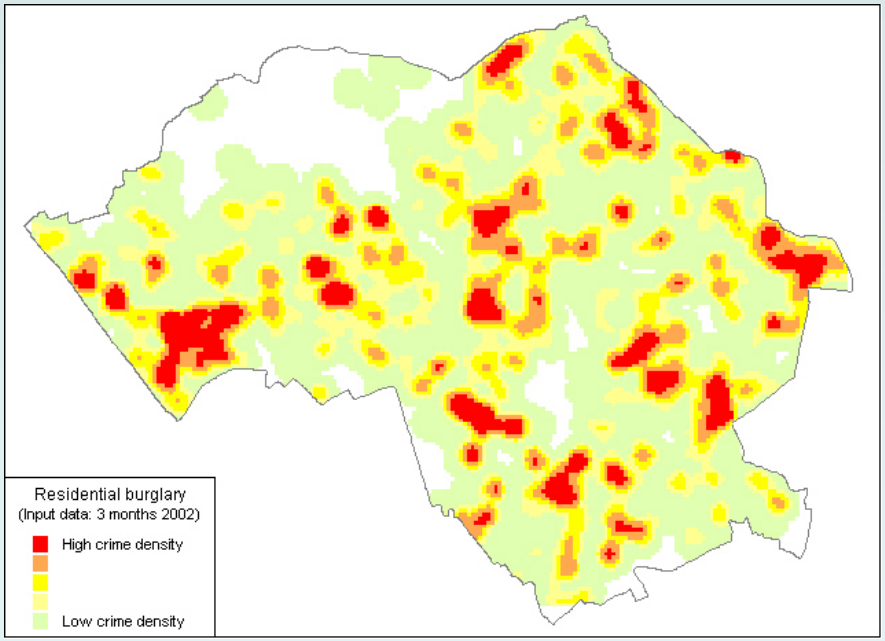
Kernel density estimation HAI: 5.8



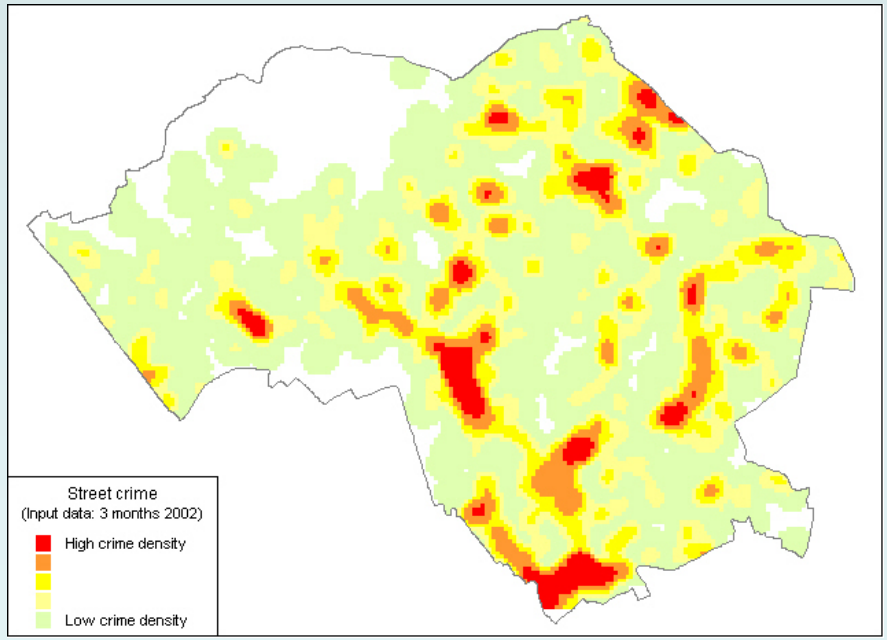
What we found out

Crime type and hotspot accuracy

Residential burglary HAI: 2.2



Street crime HAI: 5.8



UCL Jill Dando Institute of Crime Science

What we found out

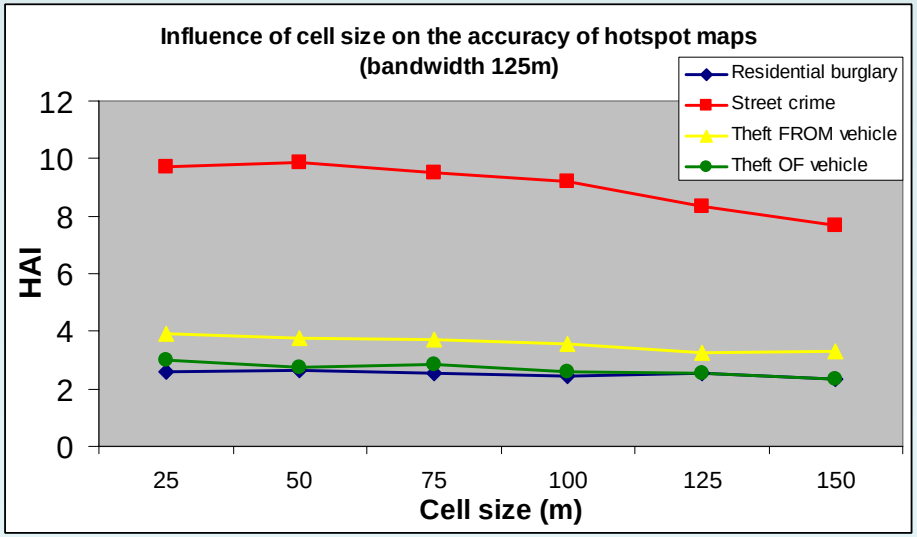
Kernel density estimation parameters and hotspot accuracy

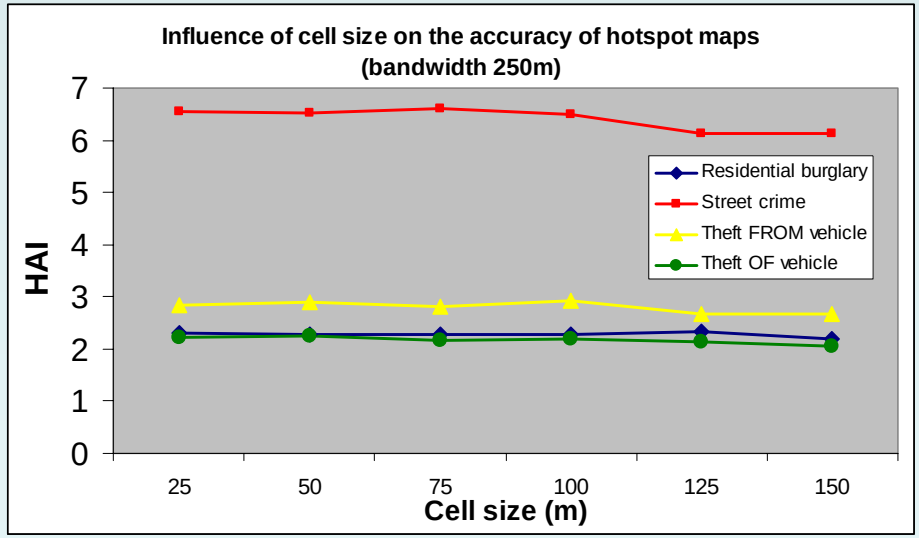
- Parameters
  - Cell size
  - Bandwidth/search radius
- Hotspot maps were being created using the Hotspot Detective defaults
- How do these parameters influence the hotspot maps accuracy?

What we found out

Kernel density estimation parameters and hotspot accuracy

- Cell size
  - Small influence on hotspot accuracy
  - Smaller cell size produces more accurate hotspot maps

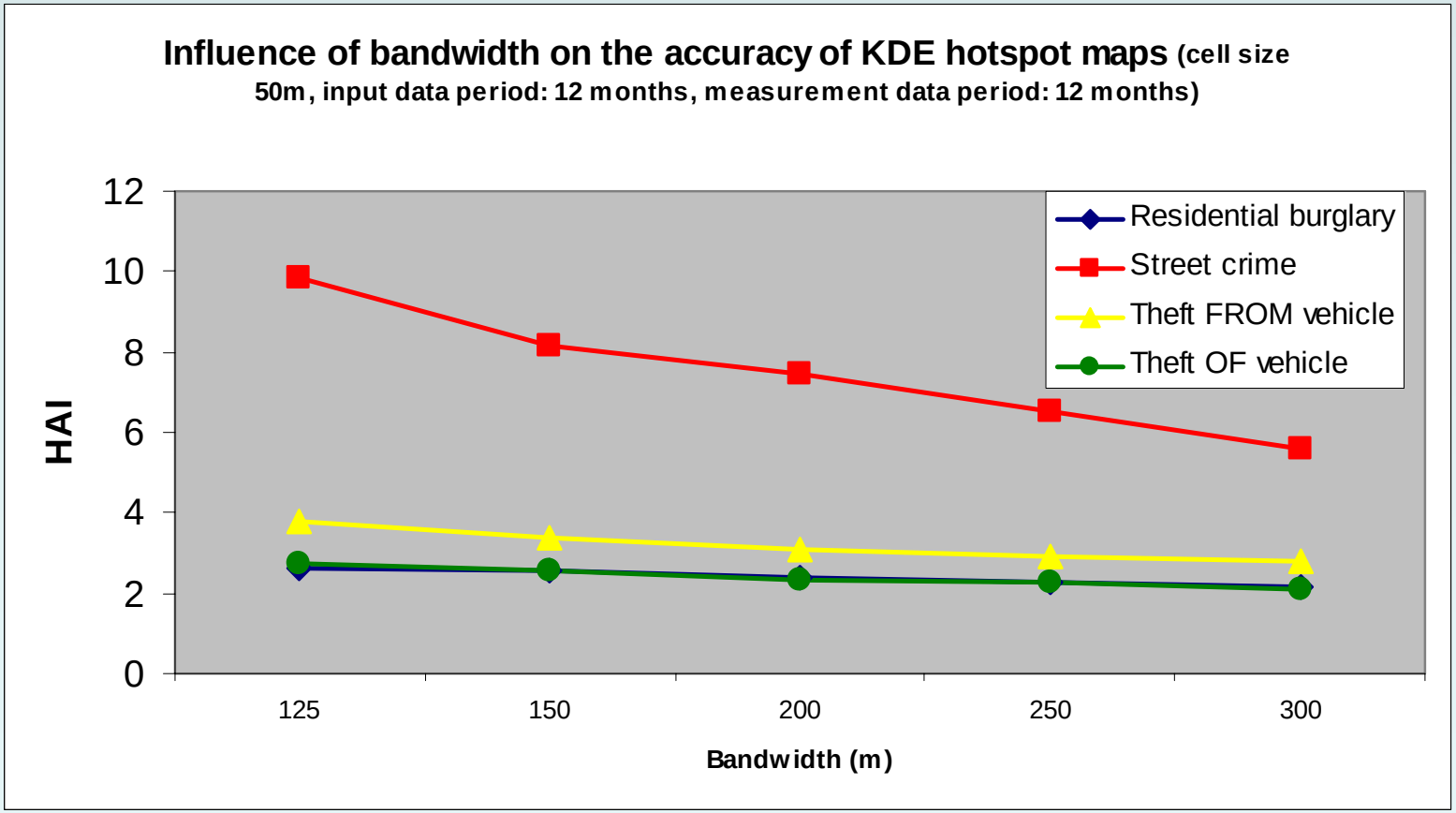




What we found out

Kernel density estimation parameters and hotspot accuracy

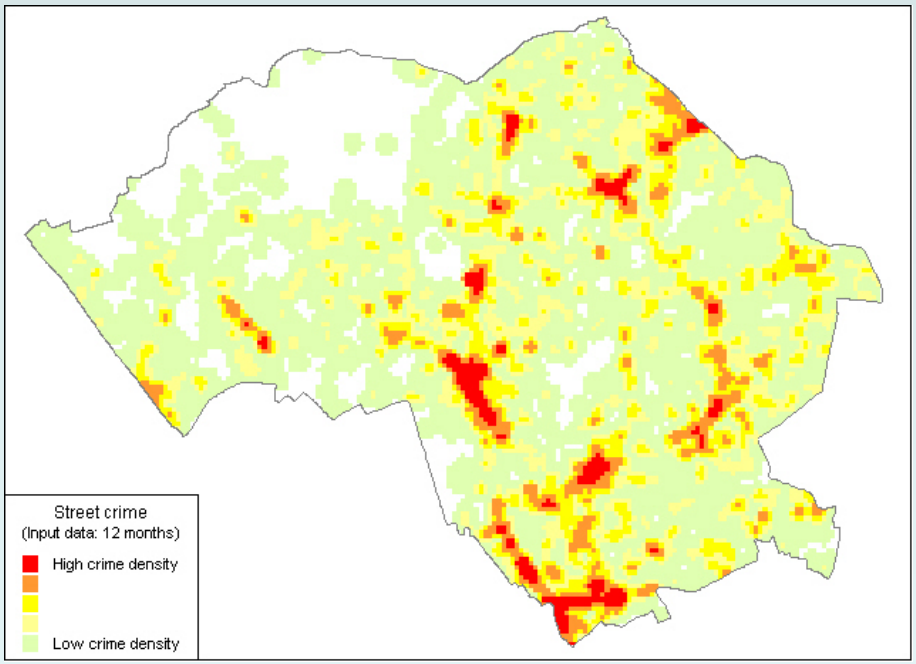
- Bandwidth
  - The smaller the bandwidth, the higher the hotspot accuracy



What we found out

Kernel density estimation parameters and hotspot accuracy

- KDE parameters



Cell size: 50m

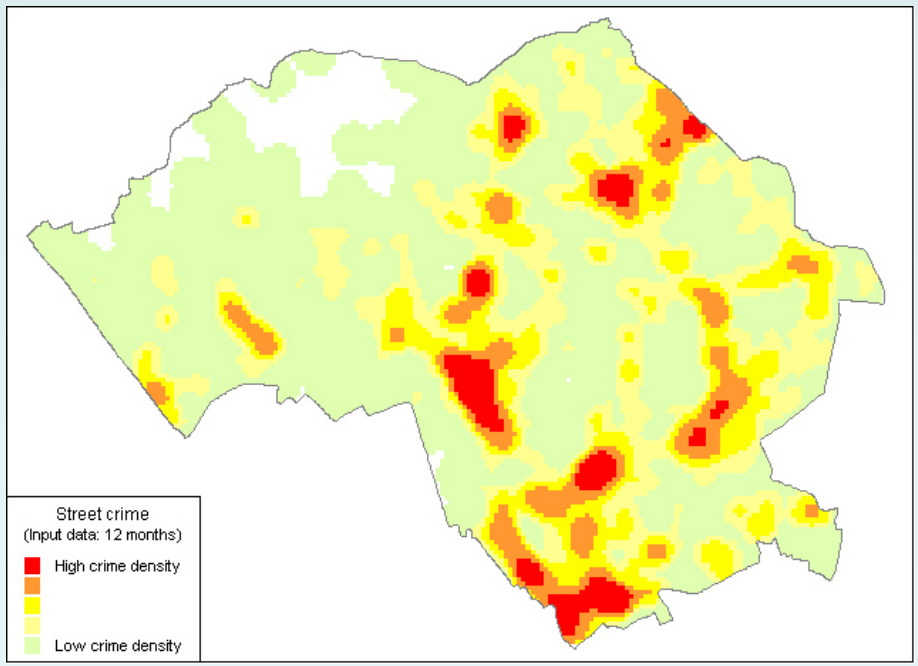
Bandwidth: 125m

HAI: 9.84

Cell size: 50m

Bandwidth: 225m

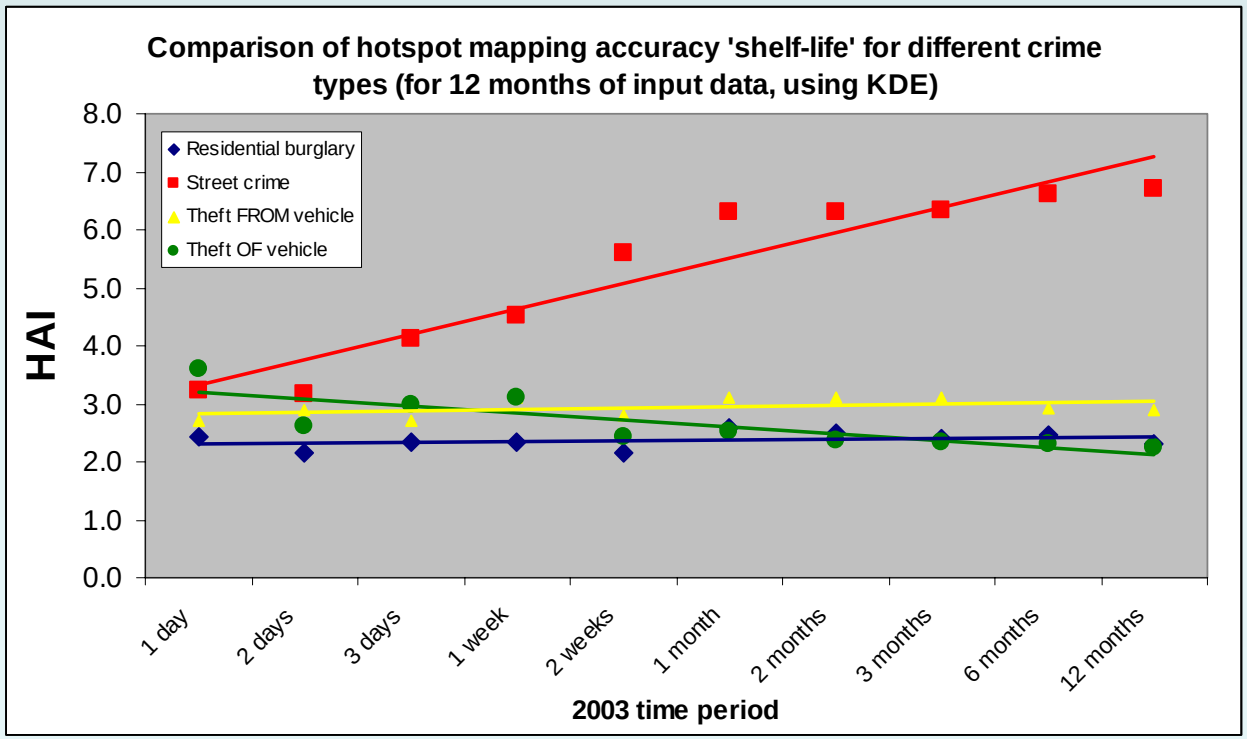
HAI: 6.51



What we found out

Shelf-life of the hotspot map

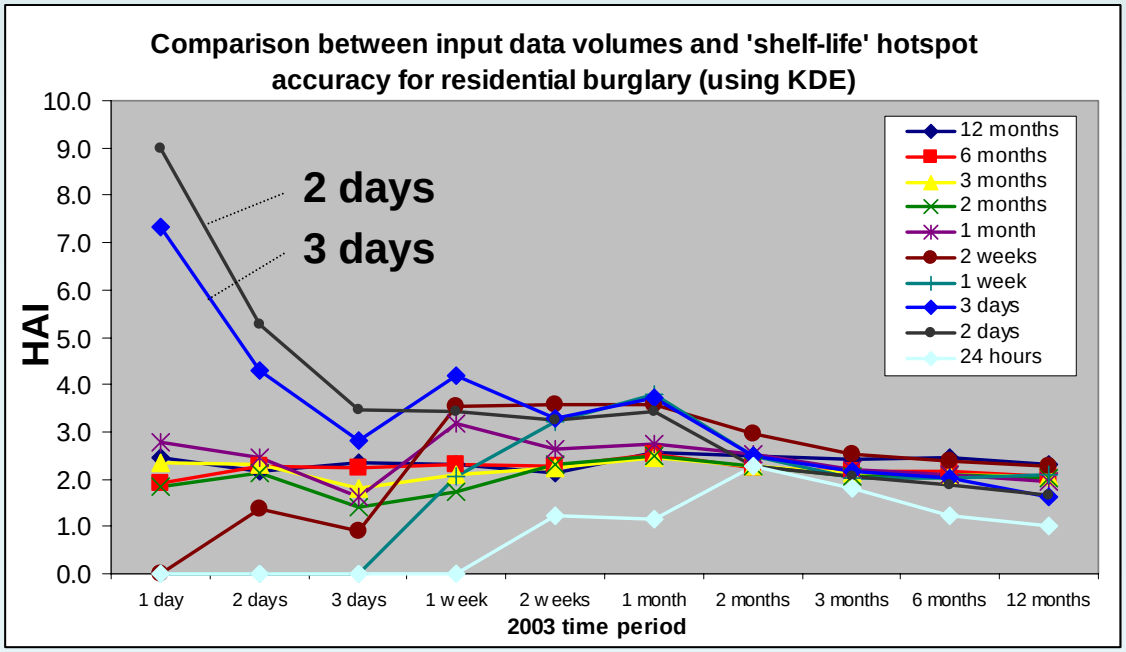
- A hotspot map's accuracy does change over its shelf-life
- Example (using 12 months of input data):
  - Street crime hotspot map accuracy improved as the map aged
  - Residential burglary and theft from vehicles hotspot maps did not deteriorate with age
  - Theft of vehicle hotspot maps showed slight deterioration over time



What we found out

Influence of input data volumes on hotspot map accuracy

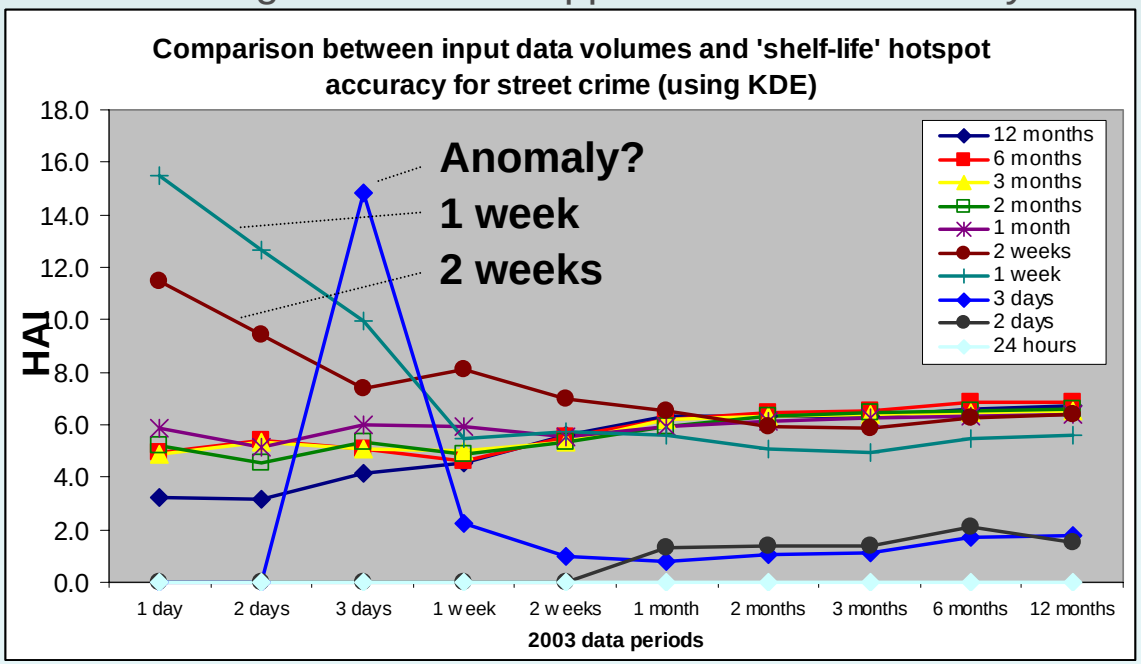
- Residential burglary
  - Influence of volume of input data became less significant at about 2 months into the maps' shelf-lives
    - i.e. just as good using last 2 days of data as it was for using 12 months
  - Hotspot maps of crime that had happened in the last 2 and 3 days were most accurate at showing what would happen in the next few days



What we found out

Influence of input data volumes on hotspot map accuracy

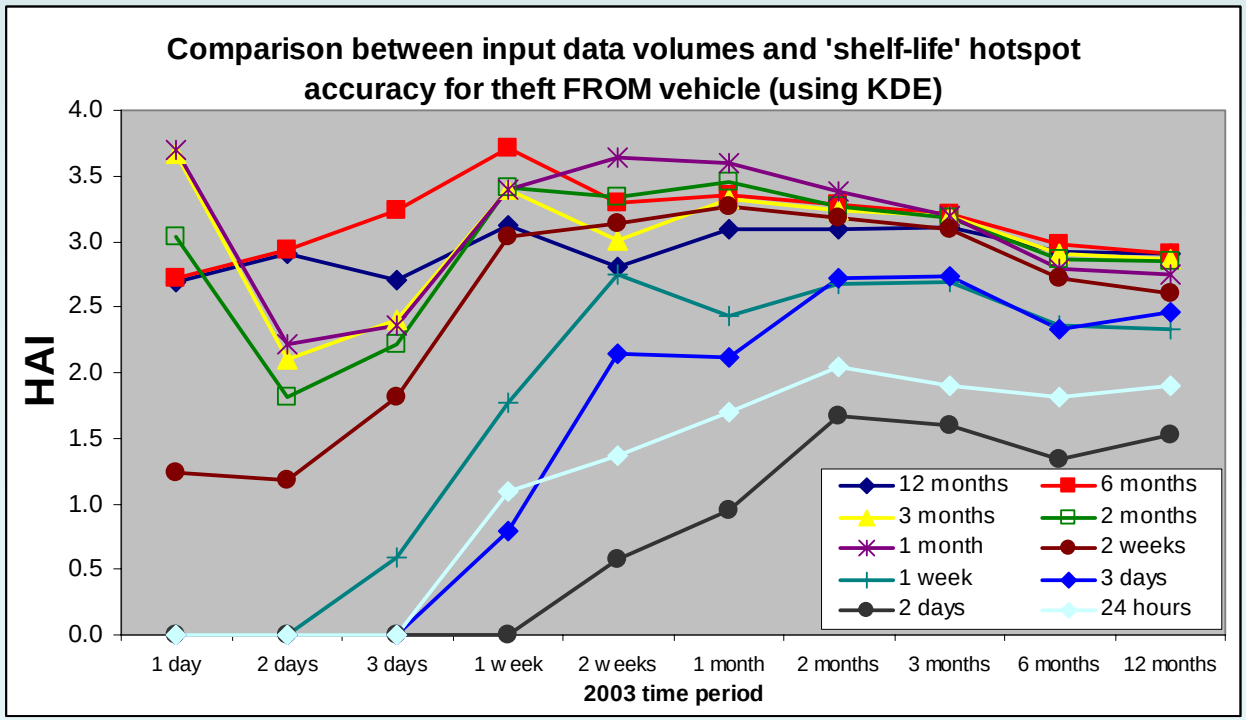
- Street crime
  - Influence of volume of input data became less significant at about 1 month into the maps' shelf-lives, apart from input data from the last few days
  - Crime data from the last few days was not accurate at showing what crimes may happen next
  - Hotspot maps of crime that had happened in the last 1 week or 2 were most accurate at showing what would happen in the next few days



What we found out

Influence of input data volumes on hotspot map accuracy

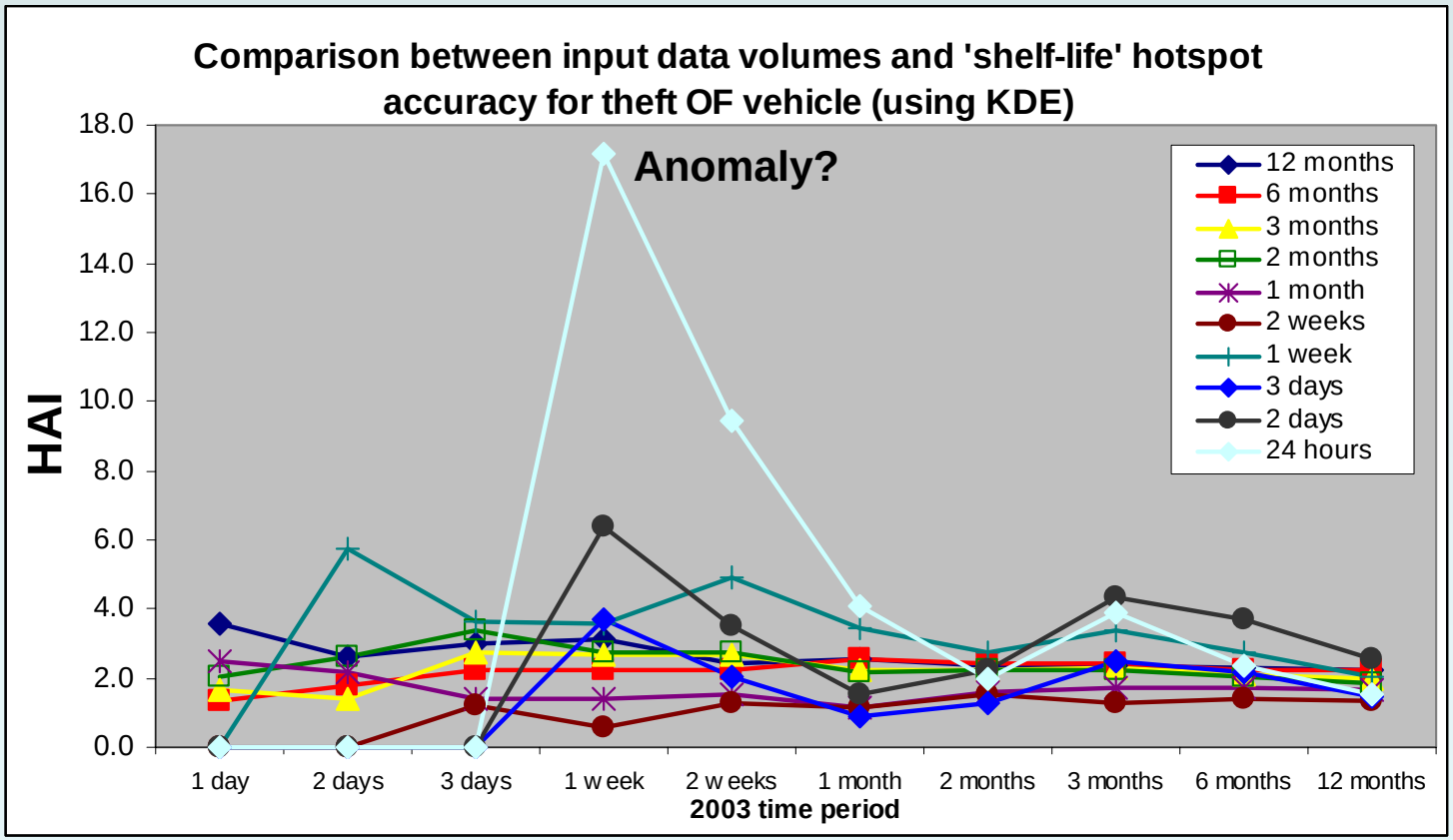
- Theft FROM vehicle
  - Input data of at least 2 weeks produced hotspot maps of consistent accuracy after a 1 month period on the shelf
  - Crime data from the last few days was not an accurate measure for showing what crimes may happen next



What we found out

Influence of input data volumes on hotspot map accuracy

- Theft OF vehicle
  - Volume of input data did not have a consistent impact on differences in hotspot accuracy for different periods into the future



Summary of findings

- Differences between crime types
  - Street crime hotspot maps are more accurate than those for residential burglary and vehicle crime
- Difference between hotspot mapping techniques
  - Kernel density estimation consistently produced high hotspot accuracy measures
  - Careful choice of parameters can optimise accuracy
- Do hotspot maps have a shelf-life/use-by date?
  - Yes, but this can differ by crime type
- How does the currency/volume of data I use influence the accuracy of my hotspot map?
  - For immediate (operational) responses: different crime types showed different data requirements
    - E.g. Residential burglary – crime data from the last 3 days produced the best results – whereas theft FROM vehicles – data from the last 6 months worked better than very current data
  - For longer term (strategic crime prevention) responses
    - Do not need long periods of historical data to produce an accurate hotspot map
    - In most cases 1 months crime data is sufficient

Considerations

Currency is important

- Both positively and negatively for identifying hotspots and predicting what happens next
  - Need to treat the production of crime hotspots by crime type, rather than running the same production line for each
  - Weighting crimes by their currency appears more complex than originally thought
    - i.e. can not just assign the same weighting formula to all crime types

What best indicates what will happen next?

- Is retrospective crime data the best indicator for targeting future actions?
- High hotspot accuracy scores for street crime appeared to be an indicator of street crime being linked to the stability of opportunities for street crime
- Mapping opportunity surfaces in space and time may be a better indicator of what happens next then just using retrospective data
  - But it is harder to do

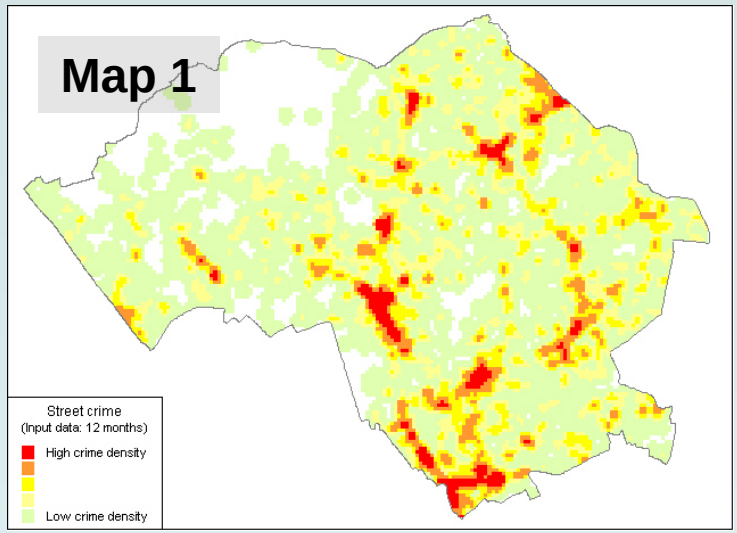
Considerations

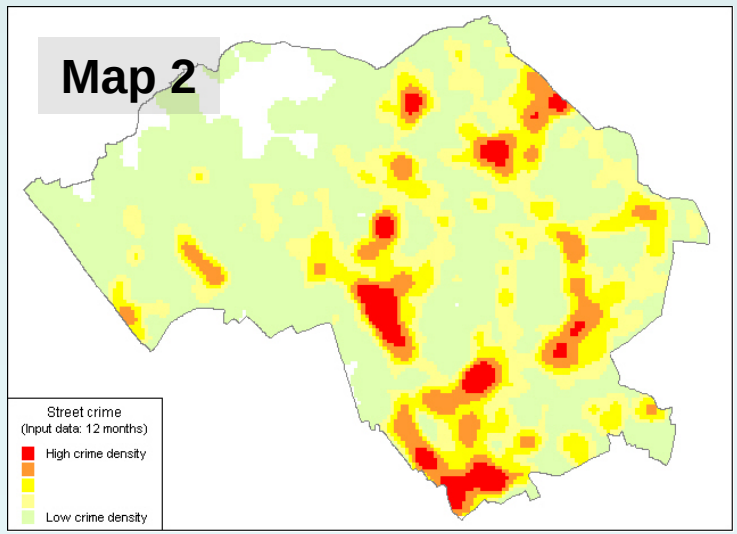
Pretty vs Practical

- Map 2 may look prettier but the top one is more accurate
- Could be more difficult to deploy resources to hotspots in Map 1 (21 different contiguous hotspots) than to Map 2 (11 hotspots)
- Measure for practicality
  - Average nearest neighbour distance between each 'hot' cell
- We do still need good cartographic design
- But does my map need to be really accurate for the purpose it will serve?

Research for other 'test dates'

- This research only used 1st January 2003
- Need to repeat analysis for other test dates and explore any differences in findings





Conclusion

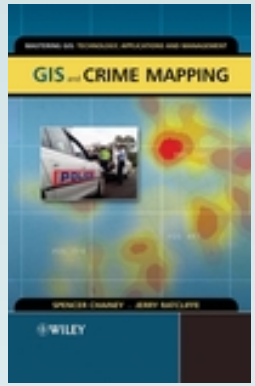
- Initiated a simple to apply methodology and some benchmark measures for hotspot accuracy
  - Shown differences between common techniques, crime types, influence of data currency/volume, and the shelf-life of hotspot maps
  - Forms a basis for others to measure and compare their hotspot map's accuracy
  - Benchmark measures for comparing other and new techniques (e.g. LISA statistics, prospective mapping, AI prediction techniques)

Thankyou

Now available in all good book shops and online book stores

‘GIS and Crime Mapping’

by Spencer Chainey and Jerry Ratcliffe.



For details see www.jdi.ucl.ac.uk

Spencer Chainey
The Jill Dando Institute of Crime Science
University College London
s.chainey@ucl.ac.uk